Method for reserving bandwidth in an Ethernet type network

The invention relates to Ethernet bus type communication networks and, more particularly, such "domestic" networks.

5

It is known that an Ethernet bus type communication network operates according to the CSMA/CD mode, in other words, by carrier sensing with collision detection. Some research work has proposed the use of priority allocation mechanisms implemented in the Ethernet packets to enable, in this

10   type of communication network, data streams to be transmitted over the Ethernet bus. However, this type of priority allocation mechanism is intended more for computer networks organized around routers. It is not suited to small communication networks of the domestic network type in which the elements are passive.

15

The object of the invention is to propose a mechanism for reserving, on at least one node of an Ethernet bus type communication network, a certain fraction of the bus bandwidth so as to enable a data stream to be transmitted, for example, an audio/video stream according to the DVB (Digital Video Broadcast) standard.

20

The idea on which the invention is based is to superimpose on the communication protocol of a digital bus, bus access arbitration by the circulation of a token between the nodes of the network so as to be able to control the times of write mode access to the bus by the nodes of the network. The term

25   token denotes a write permission which is granted to a node of the network in the sense of a so-called token-ring network. The bandwidth is divided into cycles during which the write permission of the token passes from one node to another according to a predetermined and reproducible strategy over a number of cycles.

30

The document "Rether: A Real-Time Ethernet Protocol" by Chitra VENKATRAMANI, published in November 1996, describes a network consisting of nodes having a permission to send during a time slot dependent on the previously reserved bandwidth. A control node divides up the bandwidth

35   temporally by creating a first list of nodes having requested bandwidth, and a second list of nodes not having reserved. During a cycle, the terminals of the first list send their messages first, then the terminals of the second list have their turn during the remaining time. When the end of the cycle occurs, the

message transmissions by the nodes of the second list are stopped and resume on the next cycle at the point at which they were interrupted in the second list. In this document, not all the terminals are sure to be able to have their turn during a cycle.

5

More particularly, the object of the invention is to devise a method for reserving, on at least one node of an Ethernet bus type communication network, a predetermined fraction of the bandwidth of the digital bus during a cycle; characterized in that it consists in:

10   - having a token circulate between the nodes of the network (A, B, C, D) so as to enable the nodes of the network to send in turn a data packet over the bus (1) according to a predefined sequence defining a chronological order of passage of the token between all the nodes during a cycle; and

  - in which the predetermined fraction of the bandwidth reserved for a node of

15   the network corresponds in the sequence to a certain number of occurrences of passage of the token via the node concerned.

The invention relates to digital bus type domestic communication networks comprising network nodes configured to apply the method defined above. The

20 invention ensures that each node of the network will be able to have a turn at least once during a cycle.

The method according to the invention offers the following refinements:

  - the occurrences of passage of the token via a node of the network are

25   distributed in the sequence among the occurrences of passage of the token via the other nodes of the network to avoid jitter effects;

  - the chronological order of passage of the token between the nodes of the network is defined by a master node of the network;

  - the master node, on initialization of the network, constructs a first table

30   storing, for each node of the network, information indicative of the fraction of bandwidth reserved for the node of the network and, on the basis of the first table, the master node constructs a second table storing the sequence defining the order of passage of the token between the nodes of the network.

35 The invention can be extended further to a communication device designed to be connected to a digital bus communication network, characterized in that it is configured to have a token circulate between the nodes of the network during a cycle and in that it is organized to construct a first table storing, for each node of

PF030021_PCT as filed

the network, information indicative of a fraction of the bus bandwidth reserved for the node of the network and a second table storing a sequence defining a chronological order of passage of the token between all the nodes during a cycle, the fraction of the bandwidth reserved for a node of the network

5   corresponding in the sequence to a certain number of occurrences of passage of the token by the node concerned.

Other features and advantages of the present invention will become apparent from the description of the exemplary embodiments below, taken as non-limiting

10   examples, with reference to the appended figures in which:

Figure 1 represents very schematically an Ethernet bus type domestic communication network according to an exemplary embodiment of the invention.

15

Figure 2 is a timing diagram illustrating the operation of the network nodes to apply the method according to an exemplary embodiment of the invention.

The Ethernet bus type domestic communication network illustrated in Figure 1

20   here comprises four network nodes represented by the blocks A, B, C and D, which can be consumer communication devices such as a dish antenna receiver, a digital decoder, an optical disc reader and a television. The Ethernet bus is indicated by the reference 1. It normally comprises a twisted-wire pair. Such a domestic communication network normally has a bandwidth of 100 Mb

25   per second. The Ethernet bus is one example, the present invention relates to any type of digital bus in which a token is passed between terminals, the USB bus being another example.

In this domestic communication network, Ethernet bus access arbitration via a

30   mechanism for circulating a token is superimposed on the Ethernet protocol to enable data streams to be transmitted over the network, for example audio and/or DVB video streams. The circulation of the token between the nodes A, B, C and D of the network allows the different nodes of the network write mode access to the Ethernet bus 1 in turn with no risk of collision. When it has the

35   token, a node sends over the bus 1 to a recipient node, one and only one Ethernet packet (layer 2 - OSI model), the maximum payload of which is 1500 bytes, plus, where appropriate, control packets (protocol management).

According to the present exemplary embodiment of the invention, the nodes of the network can each reserve a fraction of the bus bandwidth. In the example illustrated in Figure 1, the node A has reserved 10 Mb per second, the node B has reserved 20 Mb per second, the node D has reserved 30 Mb per second

5 and the node C has reserved 40 Mb per second. The information indicative of the fraction of bandwidth reserved by each node is collected by a master node of the network, in this case node A, and is stored in a table 1 mapped with the physical network addresses of the nodes. The physical addresses of the network nodes A, B, C and D are respectively represented by MACA, MACB,

10 MACC and MACD, and the logical network addresses of the nodes are respectively represented by ID0, ID1, ID2, ID3. These logical addresses are used in the network messages to circulate the token.

On the basis of table 1, the master node A constructs, in a table 3, in the form of

15 a list, a sequence defining the chronological order of passage of the token between all the nodes in such a way that the predetermined fraction of the bandwidth reserved for each node of the network during a cycle corresponds in the sequence to a certain number of occurrences of passage of the token by the node concerned.

20

The sequence defining the chronological order of passage of the token can be constructed by the master node A using a greater common divisor calculation between the reserved bandwidth fractions. In the example of Figure 1, the total available bandwidth is 100 Mb per second. The greater common divisor is

25 10 Mb per second. On this basis, the sequence defining the chronological order of passage of the token will include one passage occurrence for the node A, two passage occurrences for the node B, three passage occurrences for the node D and four passage occurrences for the node C. These passage occurrences are symbolized in the table 3 by the logical addresses ID0, ID1, ID2 and ID3 of the

30 nodes.

Moreover, the occurrences of passage of the token by a node of the network can advantageously be distributed evenly in the sequence among the occurrences of passage of the token by the other nodes of the network, for

35 example using the Bellman algorithm, as illustrated in table 3 in Figure 1. This distribution can be used to avoid jitter effects.

The operation of the nodes of the network to apply the method according to an

exemplary embodiment of the invention is now described with reference to Figure 2.

Initialization of the Ethernet bus and election of the master node

When the Ethernet bus is initialized, one or more nodes of the network must be configured as master nodes of the network. On receipt of an initialization network message sent from one of the nodes of the network, each node configured as a master node sends over the Ethernet bus 1 a first network message containing the physical address of the node. These first network messages are sent over the bus 1 from each node with a limited random delay. When the maximum delay for transmission of the first network message expires, the node configured as a master node which has the highest physical address is determined by default as the master node of the network. In the example of Figure 2, the master node of the network is the node A, the physical address of which is represented by MACA.

Notification of election of the master node on the network and logical numbering of the other nodes of the network

The master node of the network sends over the bus 1 a second network message to announce to the other nodes of the network that it is the master node and that it has a logical address, in this case represented by ID0.

When the second network message is received by all the other nodes of the network, each other node of the network returns over the bus 1 a third network message M3 containing the physical address of the node, for example MACB for the node B, and, optionally, information indicative of the fraction of bandwidth reserved by the node, for example 20 Mb per second for the node B. To avoid collisions of Ethernet network messages on the bus 1, the time interval between the moment of receipt of the second network message in a node and the moment of transmission of the third network message M3 is a limited random delay of maximum value T1. If a node of the network does not receive a fourth message M4 from the master node before the delay T1 counted from the moment of receipt of the second network message, it again immediately returns a third network message M3 over the bus as illustrated in the Figure 2 for the node C.

In response to the receipt of a third network message M3, the master node A returns to the sending node, for example the node B, the fourth network message M4 comprising the logical address assigned to the node, for example ID1 for the node B. The master node increments the logical addresses assigned to the nodes of the network as it receives the third network messages M3.

This process of numbering of the nodes of the network by interchanges of third and fourth network messages M3, M4 is represented by the part I of the timing diagram. At the same time that the fourth network messages M4 are sent over the bus 1, the master node A constructs the association table 1 containing, for each node of the network, the physical addresses MACA, MACB, MACC, MACD and logical addresses ID0, ID1, ID2, ID3 of the node and the information indicative of the fraction of bandwidth reserved by the node.

After a delay equivalent to two times the delay T1, counted from the moment of transmission of the second network message, the master node A constructs the table 3 containing the sequence defining the chronological order of circulation of the token between the nodes of the network as indicated above and transmits it to each node of the network.

After a delay equivalent to three times the delay T1, counted from the moment of transmission of the second network message, the nodes of the network are ready to start the token circulation mechanism illustrated by the part II of the timing diagram in Figure 2. This delay T1 is a parameter that must be set on initialization of the network according to the number of nodes connected to the bus 1 in order to optimize the time needed to initialize the network.

Each node that holds the token can write an Ethernet packet once to the bus 1, then transfers the token to a next node and so on according to the chronological order defined in the table 3. To control the circulation of the token from node to node, fifth network messages M5 and sixth network messages M6 are interchanged between the nodes of the network. In the token circulation mechanism according to an exemplary embodiment of the invention, the master node A, with logical address ID0, always has the token first and can write to the bus 1 as symbolized by the reference W in the Figure 2. It then sends the token to the next node in the sequence, which is the node D with logical address ID2 via a fifth network message M5. The fifth message M5 is an Ethernet transmission message received by all the nodes of the network to enable them

to follow in parallel the circulation of the token in their table 3. On receipt of the fifth network message, the node D returns over the bus 1 the sixth network message as an acknowledgement for the node A. This process continues thus from node to node according to the chronological order defined in the table 3.

Now, if, after a delay T2 counted from the moment of transmission of the fifth network message M5 from a node of the network, no sixth message M6 is sent over the bus, the node in question again returns the sixth network message M6 as illustrated for the node C in Figure 2. If no sixth network message M6 is sent over the bus 1 within the delay T2 as indicated above, the current node having the token, for example the node C in Figure 2, sends over the bus 1 a seventh network message M7 indicating that the next node in the sequence is no longer available, in the example, the node B with logical address ID1, and containing the logical address of the new next node in the sequence, in this case the node D with logical address ID2. In response to the receipt of the seventh network message M7, the other nodes of the network update the table 3 by deleting the occurrences of passage of the token by the failed node, in the example, the occurrences symbolized by ID1. The current node having the token then sends a fifth network message M5 containing the logical address of the new next node to continue the token circulation mechanism.

In the case where the failed node is the master node, the current node sends over the bus a network initialization message which triggers the election of a new master node and the construction of a new table 3 according to the same principle illustrated by the part I of the timing diagram.

After having sent its data, and before passing the token, a node can send to the master node a message M3 to change its bandwidth reservation. The master node returns a message M9 indicating either that the reservation has been successful or that it has failed (if, for example there is not enough available bandwidth left). If the reservation is accepted, the master node modifies its list 2. The next time the cycle passes via the master node, the latter updates the list 3 and transmits it over the network before recommencing a cycle. It should be noted that a certain percentage of the bandwidth must be kept in reserve as a safety margin to ensure that the algorithm for calculating the list for the passage of the token is executed correctly, given the approximations made to map the bandwidth reservation with the number of occurrences of the token. As an example, this percentage may be approximately 10%.

PF030021_PCT as filed

When the token returns to the first occurrence of the master node in the table 3, the master node A recommences the cycle illustrated by part II of the timing diagram.

It is possible to provide for the master node A, before recommencing the cycle illustrated by part II of the timing diagram, to cyclically run a procedure to update the tables 2 and 3 to take account of the changes of configuration of the network, in particular the connection of new peripheral devices. This update procedure can be carried out with a period T3 very much longer than the delays T1 and T2. According to this update procedure, the master node A sends over the bus an eighth network message M8 enabling the new devices connected to the bus to be identified as a new node of the network as illustrated by the reference E in Figure 2. Within the delay T2 counted from the moment of transmission of the eighth network message M8, the master node A then waits for the receipt of one or more third network messages M3. Figure 2 represents, by way of example, the exchange of a third network message M3 and a fourth network message M4 between the master node A and the new node C. Each time a message M3 is received, the time counter is reset. At the end of the delay T2, the master node reconstructs, if there has been a change, a new table 3 and transmits it over the bus 1.

In practice, for the token circulation mechanism to operate correctly, the delay T1 is set to be greater than the delay T2 and less than the delay T3. As an example, T1 can be set to 100 ms, T2 to 5 ms and T3 to 3 minutes.

With this token circulation mechanism, it is possible to guarantee each node of the network the use of a certain fraction of the available bandwidth on the bus for the transmission of data streams.

This token circulation mechanism can be implemented simply in the nodes of the network by software configuration.